August 18, 1993 March 15,1990 - July 31, 1992

Modified aequorin shows increased bioluminescence activity.

N00014-90-J-1704

Douglas C Prasher

Biology Department
Woods Hole Oceanographic Institution
Woods Hole, MA 02543

Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217-5000

Current address: USDA, APHIS Ph: 508-563-9303
Bldg #1398
Otis ANG Base, MA 02542

Distribution Unlimited

Aequorin belongs to a unique class of photoproteins that emit light upon the binding of certain metals, calcium being the principal intracellular activator. This 'reporting function' of the metal-binding is instanteous and is very easy to quantitate experimentally. The project objective was to develop a variety of recombinant forms of aequorin so they can be employed as metal biosensors. Three calcium-binding sites of aequorin were modified to examine their roles in the calcium-dependent luminescence as well as potentially binding other metal ions. Aequorins having Site 2 substitutions unexpectedly produce more light than wild-type aequorin.

calcium, bioluminescence, photoprotein, mutagenesis

Unclassified Unclassified Unclassified UL

FINAL PROGRESS REPORT

Modified Aequorin Shows Increased Bioluminescence Activity

15 March 1990 thru 31 July 1992

INTRODUCTION

Aequorin belongs to a unique class of photoproteins that emit light upon the binding of certain metals, calcium being the principal intracellular activator. This 'reporting function' of the metal-binding is instantaneous and is very easy to quantitate experimentally. The project objective was to develop a variety of recombinant forms of aequorin so they can be employed as metal biosensors.

Since aequorin is a member of the calcium-modulated protein superfamily, the amino acid sequences in the calcium-binding sites of aequorin are analogous to those of numerous other eucaryotic calcium-binding proteins. These proteins bind μM Ca^{2+} in the presence of physiological concentrations of Mg^{2+}. The binding of calcium to the polypeptide is assumed to induce a conformational change resulting in the oxidation of the bound luciferin with the production of light. Because the metal ion is not believed to be involved in the catalytic oxidation, substitution of the codons encoding the amino acids assumed to serve as metal ligands can lead to an aequorin polypeptide having an altered metal-binding specificity while potentially retaining the luminescent reporting function.

The **specific aims** of the project were to (1) examine the role of the three calcium-binding sites of aequorin in its luminescence, (2) examine the role certain calcium-ligands have on the binding of calcium and other metals, (3) engineer the polypeptide such that the metal specificity of the luminescence favors other alkaline earth metals or transition metals, and (4) produce recombinant aequorin containing selenomethionine for X-ray structural analysis by a collaborator Dr. Linda Hannick (NRL).

The funding period was terminated prematurely due to the Principal Investigator leaving the Woods Hole Oceanographic Institution. The grant was originally scheduled to end during March 1993.

ACCOMPLISHMENTS

Recombinant Aequorin Encoded by Three Wild-Type cDNAs

Recombinant forms encoded by the three wild-type aequorin cDNAs, *AEQ1*, *AEQ2*, and *AEQ3*[1] were made available during the grant period. This was the first time that purified aequorin isotypes were available for comparison[2]. The gene products of *AEQ1* and *AEQ3*, recombinant AEQ1 and AEQ3 (rAEQ1, rAEQ3), were available prior to the grant period. An *AEQ2* expression vector was produced early in the grant period. Recombinant AEQ1 differs from rAEQ2 and rAEQ3 at 19 and 16 amino acid residues, respectively. Following charging with synthetic luciferin, their total light yield differed by a maximum of 49% (Table 1). However, while rAEQ2 has the highest light yield of the three recombinant aequorins. it was also the most difficult to purify since it was produced in *E.coli* at a substantially lower level than rAEQ1 and rAEQ3. Since rAEQ2 and rAEQ3 differ by only three residues and these

[1]Prasher *et al.* (1987), *Biochemistry 26*:1326.

[2]The isoaequorins, purified from native aequorin preparations, described by Shimomura (*Biochem. J. 234*:271,1986) are not necessarily pure.

aequorins show the greatest difference in the light yield, it is assumed the amino acid replacements at positions 70,71, and 164 (Table 1) are primarily responsible for the decreased light yield in rAEQ1 and rAEQ3 when compared to rAEQ2.

Ca^{2+}-Binding Site	Amino Acid Position	rAEQ1.3	rAEQ3.2	rAEQ2.1
		1.95 E16 hv/mg 51%	2.56 E16 hv/mg 67%	3.84 E16 hv/mg 100%
	12	Pro	Ser	Ser
1	37	Arg	Lys	Lys
	70	Glu	Glu	Gly
	71	Ala	Ala	Asp
	85	Glu	Asp	Asp
	88	Glu	Ala	Ala
	95	Arg	Lys	Lys
	98	Ser	Thr	Thr
	99	Glu	Asp	Asp
	102	Lys	Glu	Glu
	103	Arg	Lys	Lys
	105	Ser	Ala	Ala
	108	Gln	Glu	Glu
	109	Ile	Pro	Pro
	114	Leu	Ile	Ile
	123	Ile	Val	Val
2	132	Ser	Thr	Thr
	142	Ser	Ala	Ala
3	164	Ser	Ser	Asn

Table 1. Relative activities of recombinant aequorins encoded by three different cDNAs. The amino acid residues at each heterogenous position is listed for each recombinant aequorin.

Site-Directed Mutagenesis of *AEQ3*.

A variety of aequorin expression plasmids having nucleotide replacements in the coding regions were constructed to address the role each Ca^{2+}-binding site plays in the luminescent reaction. These plasmids were variations of pAEQ3.1, the expression plasmid containing the *AEQ3* cDNA (Fig. 1). Production of apoaequorin in *E.coli* is controlled by transcription initiation at the lambda p_L promoter. Encoded on a separate but compatible plasmid is a temperature-sensitive form of the lambda cI repressor. At the permissible temperature (30°C), cI binds at the p_L promoter inhibiting transcription. Apoaequorin synthesis can be induced by a thermal shock at 42°C which inactivates the cI repressor.

The site-directed mutagenesis method of Lewis and Thompson (*NAR 18*:3439,1990) was used to generate the altered aequorin expression plasmids. "The system is based on the use of a mutagenic oligonucleotide to confer antibiotic resistance to the mutant DNA strand. The system employs a phagemid vector, pSELECT-1, which contains two genes for antibiotic resistance. One of these genes, for tetracycline resistance, is always functional. The other, for ampicillin resistance, has been inactivated. An oligonucleotide is used that restores ampicillin resistance to the mutant strand during the mutagensis reaction. This oligonucleotide is annealed to the single-stranded DNA (ssDNA) template at the same time as the mutagenic oligonucleotide and subsequent synthesis and ligation of the mutant strand links the two. The DNA is transformed into a repair minus strain of *E.coli* (BMH 71-18) and the cells are grown in the presence of ampicillin, yielding large numbers of colonies. A second round of transformation in JM109 ensures proper segration of mutant and wild-type plamsids and results in a high proportion of mutants."[3]

[3]Altered Sites in vitro Mutagenesis System, Technical Manual, Promega Corp.

The *Hin*dIII-*Bam*HI fragment in the aequorin coding region in pAEQ3.1 was removed and subcloned into pSELECT-1 for mutagenesis. After mutagenesis, it was to be subcloned back into the remaining portion (mostly vector) of pAEQ3.1. At the beginning of the funding period, we found the efficiency of the second subcloning step to be extremely low. We speculated the low effeciency was due to the presence of a GC region 3' to the *Bam*HI site in pAEQ3.1 which remained from the original cloning of the aequorin cDNAs[4]. This unstable DNA was removed first by designing oligonucleotide primers that flank the GC region, and second, using them in the polymerase chain reaction to amplify the entire plasmid less the GC region. The resulting plasmid is called pAEQ3.2.

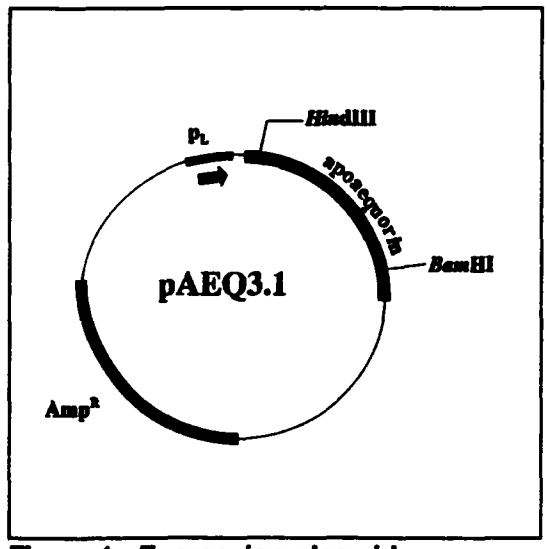

Figure 1. Expression plasmid containing the inducible apoaequorin cDNA.

A second problem appeared when the DNA sequence was confirmed following the first time the mutagenesis procedure was used. Codon-157 was discovered to code for an isoleucine instead of an arginine as originally identified in the *AEQ3* cDNA[1]. This codon was repaired prior to the construction of any additional aequorin expression vectors.

We found, not unexpectedly, the efficiency of the mutagenesis procedure was sequence dependent. It varied from 3.6% to 50% depending upon the mutant oligonucleotide used.

Calcium-Binding Sites in Aequorin.

Aequorin is believed to possess three calcium-binding sites based on its sensitivity to calcium and amino acid sequence alignments with other calcium-binding proteins. Figure 2 shows the amino acid sequences of the three sites. When bound, Ca^{2+} is predicted to be coordinated by six residues whose positions in space approximate that of the vertices of an octahedron. Five of the ligands (X,Y,Z,-X,-Z) have oxygen-containing side chains where the sixth ligand (-Y) is the carbonyl oxygen in the peptide bond and can be supplied by any amino acid. The ligand X is always Asp (D) and ligand -Z is generally Glu (E). The Gly at position 6 is conserved and the residue at position 8 is Ile, Val, or Leu in calcium-binding proteins.

Aequorin Mutants

A variety of aequorin mutants were generated that were designed to examine the role each of the three calcium-binding sites play in the luminescent reaction. Four 'types' of amino

Ligand:	X	Y	Z	-Y	-X	-Z
Site 1 Sequence:	**Asp**	Val	**Asn**	His	**Asn**	Gly	Lys	Ile	**Ser**	Leu	**Asp**	**Glu**
Site 2 Sequence:	**Asp**	Lys	**Asp**	Gln	**Asn**	Gly	Ala	Ile	**Thr**	Leu	**Asp**	**Glu**
Site 3 Sequence:	**Asp**	Ile	**Asp**	Glu	**Ser**	Gly	Gln	Leu	**Asp**	Val	**Asp**	**Glu**

Figure 2. Alignment of the amino acid sequences of the Ca^{2+}-binding sites in aequorin. The putative ligand residues are in bold.

[4]Prasher *et al.* (1985), *BBRC 126*:1259.

acid replacements designed to disrupt Ca^{2+}-binding were effected in each 12-residue site. These include (1) the first Asp was replaced with a Ser residue, (2) the last Glu was replaced with a Ser residue, (3) the first Asp and last Glu were both replaced with Ser residues, and finally, (4) the conserved Gly in position 6 was replaced with an Ala residue.

Amino Acid Substitution	Ca^{2+}-Binding Site Modified	Expression Plasmid	Relative Activity
D31S	1	pAEQ3.13.5, .6	44%
E42S	1	pAEQ3.12.3,.7,.8	94%
D31S+E42S	1	pAEQ3.32.1	
G36A	1	pAEQ3.14.1,.2	73%
D124S	2	pAEQ3.20.3,.6	231%
E135S	2	pAEQ3.22.3,.6	129%
D124S+E135S	2	pAEQ3.34.1	
G129A	2	pAEQ3.15.8	142%
G129R	2	pAEQ3.28.7,.10	
D160S	3	pAEQ3.18.1,.4	
E171S	3	pAEQ3.24.1	
D160S+E171S	3	na	
G165A	3	pAEQ3.16.3,.5,.6	
N33D+H34G	1	pAEQ3.26.6,.9	
D124H+D126C+ N128H+E135H	2	pAEQ3.38.14,.21, .26	

na= not available

Table 2. Expression plasmids encoding mutant aequorins generated during the funding period.

Total light yields from six of these recombinant aequorins are listed in Table 2 and shown graphically in Figure 3. Although expression plasmids have been constructed for all, purified recombinant protein encoded by each was not available at the end of the funding period. Of those mutant aequorins compared, the mutants having amino acid replacements in Site 1 show reduced activity while those in Site 2 show enhanced activity when compared to wild-type aequorin (ie that encoded by pAEQ3.2). The enhanced activities of the Site 2 mutants contrast with previously published results[5].

Since aequorin mutants having a substitution for each Gly described by Tsuji *et al.*(1986) are different (G→R) from ours (G→A), direct comparison is difficult. Hence, we constructed the expression vector encoding the G129R mutant for direct comparison. Unfortunately, purified mutant protein was not available at the end of the funding period.

The expression plasmid encoding the mutant aequorin having Gly-165 replaced with Ala (G165A) was especially difficult to construct and could only be produced with additional modifications. The correct nucleotide substitution was made in the pSELECT derivative during the mutagenesis procedure. However, upon subcloning the modified *Hin*dIII-*Bam*HI fragment into the expression vector fragment, the recombinant expression plasmid (pAEQ3.16) always contained an additonal Gly codon inserted between those coding for Trp-180 and Tyr-181. Even though the remainder of the DNA sequence in

[5]Tsuji *et al.*(1986) *Proc.Natl.Acad.Sci.USA 83*:8107-11.

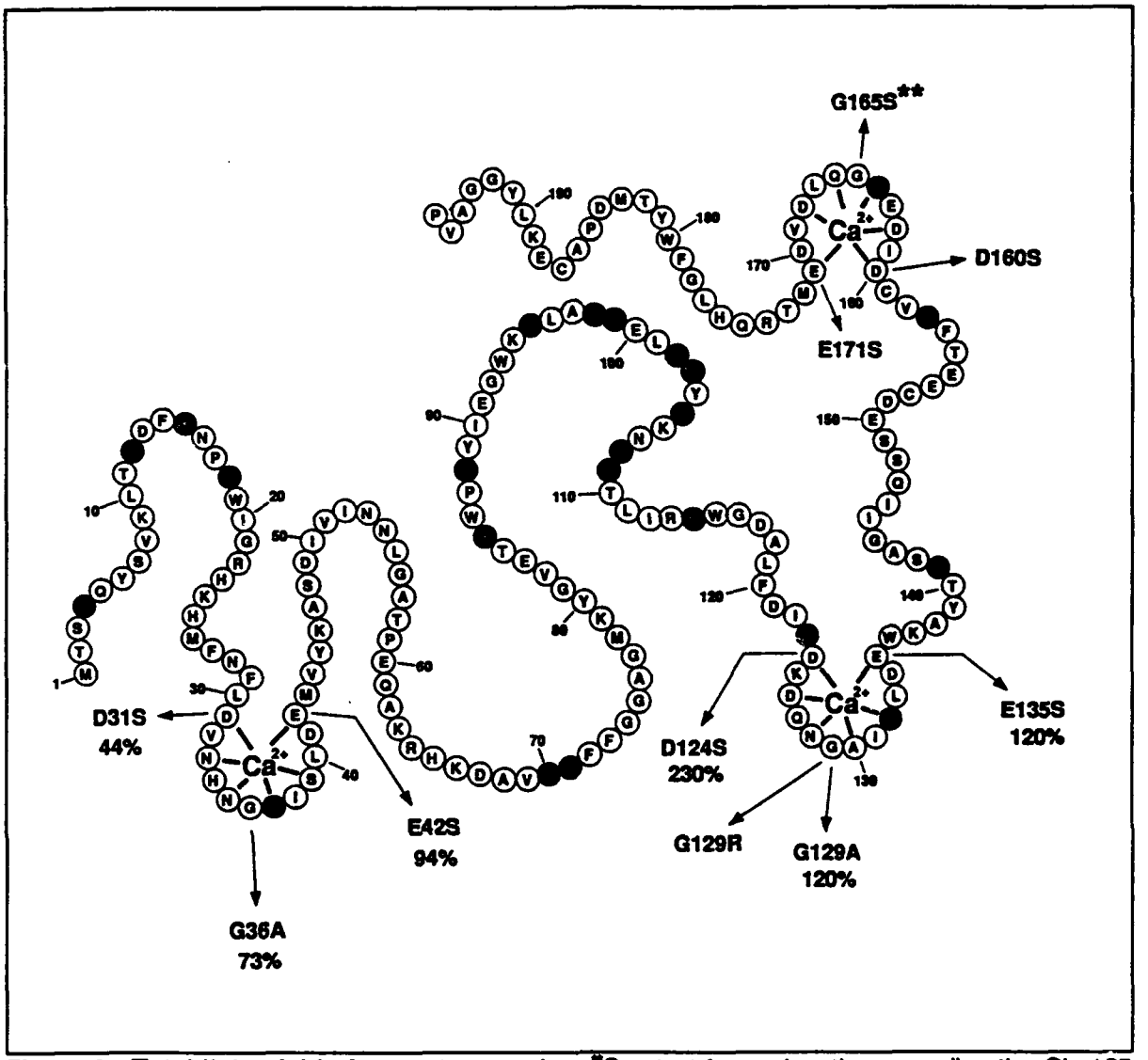

Figure 3. Total light yield of mutant aequorins. **See text for explanation regarding the Gly-165 replacement. The shaded residues represent those positions which show heterogeneity[1].

the expression vector was correct, the gene product could not be produced in *E.coli* to levels where it could be observed on a SDS PAGE gel. Since this result occurred in two independent experiments, we suspect the insertion of the Gly codon may act as a second-site suppressor for some unknown enzymatic activity associated with apoaequorin. It may accomplish this by causing the mutant polypeptide to be extremely prone to proteolysis in the *E.coli* host. However, even when a *lon*⁻ host (*ie.*protease deficient, PAM163) was used, production of apoaequorin was not observed.

Expression vectors encoding aequorin double-mutants also were generated to address the role of each Ca^{2+}-binding site. Both Asp-1 and Glu-12 in each of the three Ca^{2+}-binding sites were replaced with Ser (Table 2). Unfortunately, purified preparations of the recombinant proteins encoded by these expression vectors were not available for comparison purposes at the end of the funding period.

The light emission from aequorin is known to be affected by Mg^{2+} which acts as an antagonist to the Ca^{2+}-dependent luminescence. Analysis of the amino acid sequences of each of the three Ca^{2+}-binding

sites in aequorin according to the 'rules' described by Collins *et al.* (1991)[6] suggested to us that Site 1 is much more like a Ca^{2+}-Mg^{2+} site than Sites 2 or 3. To test this hypothesis, a mutant aequorin was generated with Site 1 having a sequence more like Ca^{2+}-specific sites. The Asn-33 residue (Y ligand) was replaced with an Asp and His-34 was replaced with a Gly (N33D+H34G). As before, the expression vector was available but the recombinant mutant protein was not available at the end of the funding period.

Work had started on producing aequorins which favor the binding of Cu^{2+} or Zn^{2+} at Site 2. These two cations are utilized by several other metalloproteins in which the preferred ligand groups are primarily imidazole nitrogens and cysteine sulfurs. Prior to the grant period, molecular modelling studies (Dr. Pattbiraman, NRL) showed that Site 2 in aequorin was analogous to Site III in calmodulin. The amino acid sequence of Site 2 could be modelled using the coordinates of Site III in calmodulin. The predicted distances between Ca^{2+} and its ligands were very reasonable (ie. 2.1-2.5 A).

Modelling of Site 2 having ligands replaced with histidines and cysteines showed that one configuration (D124H, D126C, N128H, E135H) had predicted bond distances from the ligand atoms to the metal ion sufficient to bind Cu^{2+} or Zn^{2+}. The expression vector encoding this aequorin (pAEQ3.38, Table 2), but not purified protein, was available at the end of the funding period.

Purification of Aequorin Mutant Proteins

The *E.coli* host strains containing the expression plasmids were grown under conditions summarized above. The cell pastes were always analyzed via SDS PAGE for over-production of the recombinant apoaequorin.

The following scheme was developed to purify the aequorins from an *E.coli* cell paste. Cells were lysed using lysozyme followed by several freeze-thaws. After a DNase treatment and several Triton washes of the particulate fraction, the recombinant protein was solubilized using a modification of the method described by Frankel *et al.* (1991)[7]. Following an extensive dialysis to remove much of the sarkosyl, the preparation was fractionated first, via gel filtration (Sephacryl S-200) and, second, via ion-exchange chromatography (Fractogel DMAE, EM Separations). A large number of extracts at various stages of purification were available at the end of the funding period. The extracts are summarized in Table 3.

Aequorin Designation	Amino Acid Replacement	Total mg Protein Partially Purified	Total mg Protein Purified
rAEQ1.3	wt	110mg	30mg
rAEQ2.1	wt	15.8	-
rAEQ3.2	wt	300	130
rAEQ3.12	E42S	16.3	-
rAEQ3.13	D31S	2.63	-
rAEQ3.14	G36A	19.9	-
rAEQ3.15	G129A	4.32	-
rAEQ3.18	D160S	8.9	-
rAEQ3.20	D124S	1.2	-
rAEQ3.22	E135S	1.28	-
rAEQ3.24	E171S	0.43	-

Table 3. Aequorin proteins available at the end of the funding period.

[6] *Biochemistry 30*:702-707 (1991)

[7] Frankel, S, R Sohn, & L Leinwand (1991) *Proc. Natl. Acad. Sci. USA 88*:1192-1196.

Selenomethionyl-Labelled Apoaequorin

During the funding period, we were cooperating with Linda Hannick (NRL) on the three-dimensional structure of aequorin. Thus, we supplied the NRL with several mg of purified rAEQ3.2 but we also worked towards the production of apoaequorin having incorporated selenomethionine instead of methionine. Such a recombinant protein could be very useful for analysis by multiwavelength anomalous diffraction[8]. The methionine auxotroph DL41[7], originally suggested to us by John Horton (Columbia Univ), was shown not to be a suitable host for the aequorin expression system for reasons unknown. We found the alternative strain PAM163[9] (*met lon*) to be a far superior host for the production of apoaequorin grown in the presence of selenomethionine (determined via SDS PAGE). The decreased protease activity (*lon*) in this strain probably accounts for the higher levels of apoaequorin production.

The cell pastes have been transferred to the NRL.

Benzyl-Luciferin (Subcontract with Connecticut College)

Bruce Branchini in the Chemistry Department synthesized several batches of benzyl-luciferin for this project. The benzyl derivative is much easier to synthesize than native coelenterazine (Milt Cormier, personnal communication). The quality of the batches of highest purity were analyzed by their ability to 'charge' apoaequorin. Each of the wild-type recombinant apoaequorins charged with luciferin of Batch II produced higher light yields, while batches IIIB and IVA were 10^2-10^3 less effective. As a consequence, the batches of poor quality were used for routine analysis (eg. assaying column fractions during aequorin purification).

Batch #	Purity[1]	Total mg Protein	Relative Activity[2]
I	~80%	3mg	
II	91%	8mg	1.00
IIIB	97%	75mg	0.012
IVA	98%	46mg	0.006
IVB	75%	12mg	

[1]Thin layer chromatography performed by B. Branchini.
[2]Tested by charging rAEQ1, rAEQ2, and rAEQ3 and measuring total light yield.

Table 4. Analyses of benzyl-luciferin preparations.

[8]Hendrickson, WA, JR Horton, & DM LeMaster (1990) *EMBO J 9:*1665-1672.

[9]Johnson (1977) *Genet.Res.30:*273.

PUBLICATIONS RESULTING FROM THIS FUNDING

1993 L.I. Hannick, D.C. Prasher, L.W. Schultz, J.R. Deschamps, and K.B. Ward. Preparation and Initial Characterization of Crystals of the Photoprotein Aequorin from *Aequorea victoria*. **Proteins, Structure, & Genetics** 15: 103-107.

RELATED PUBLICATIONS TO THIS WORK

1993 Chris W. Cody, Douglas C. Prasher, William M. Westler, Frank G. Prendergast, and William W. Ward. Chemical Structure of the Hexapeptide Chromophore of the *Aequorea* Green-Fluorescent Protein. **Biochemistry** 32: 1212-1218.

1992 Dennis J. O'Kane, and Douglas C. Prasher, Evolutionary Origins of Bacterial Bioluminescence. **Molecular Microbiology** 6(4):443-449.

1992 Douglas C. Prasher, Virginia K. Eckenrode, William W. Ward, Frank G. Prendergast, and Milton J. Cormier. Primary structure of the *Aequorea victoria* green-fluorescent protein. **Gene** 111(2):229-233.

PATENTS

U.S. Patent Application **"Modified Apoaequorin Having Increased Bioluminescent Activity"**. Filed December 1, 1992.